Intra-industry firm heterogeneity, myopic adaptation and exit hazard: A fitness landscape approach to firm survival and learning[1]

Eshref Trushin[a] and Mehmet Ugur[b]

[a] *de Montfort University, Department of Economics;* [b] *University of Greenwich Business School*

Abstract

We draw on insights from the fitness landscape literature and from models of firm dynamics with learning to hypothesize that: (i) firms in industries with higher company age or size heterogeneity have higher exit hazard after controlling for age, size, and a variety of other predictors of firm survival; and (ii) higher levels of R&D investment mitigate the hazard-increasing effects of industry firm heterogeneity after controlling for the direct effects of R&D intensities at industry and firm level. We test for these novel sources of selection with evidence from a panel dataset of 35,136 R&D-active UK firms from 1998 to 2012 and a range of discrete-time hazard estimators. The findings, which remain robust to multiple sensitivity checks, offer two novel contributions to the literature: (i) firm heterogeneity is not just a passive precondition for subsequent selection process in industry evolution; this heterogeneity enhances selection as more firms might be stranded in suboptimal positions; (ii) firms in more heterogenous industries can mitigate the hazard-increasing effects through R&D investment that facilitates adaptation and search for better fitness locations.

Key words: Company survival, rugged fitness landscape, firm heterogeneity, R&D, NK model

M.Ugur ORCID: 0000-0003-3891-3641; E.Trushin ORCID: 0000-0001-9058-4262

[1] This work was supported by the ESRC under grant ES/K004824/1. We thank the data host, the UK Data Service, for their help and support in accessing the data. The standard disclaimer of the data host applies: "the use of these data does not imply the endorsement of the data owner or the UK Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce national statistics aggregates". We also thank Edna Solomon for contributions towards compilation of the dataset used in this analysis. The views expressed here are those of the authors, who are responsible for any errors or omissions.

Intra-industry firm heterogeneity, myopic adaptation and exit hazard:

A fitness landscape approach to firm survival and learning

1. Introduction.

Significant and persistent heterogeneity of firm characteristics within an industry is well documented and puzzling (Dosi *et al.*, 1997), and research that pays attention to firm heterogeneity now spans several areas in growth theory, international economics, organization studies, and industrial organization. In these areas, the entry and exit of heterogenous firms is the basis of industry evolution with creative destruction (Castellaci, 2011), however, the question of how heterogeneity itself affects firm strategies and survival is usually overlooked.

The aim of this paper is to address this gap and offer two contributions to firm survival modelling and estimation. First, we argue that firm age or size heterogeneity in an industry is an indicator of fitness landscape ruggedness, which increases selection and churning independently of direct effects of firm age or size. We mainly infer the causal relationship between age/size heterogeneity and exit hazard from the fitness landscape literature, which suggests that rugged (multi-peak) industry fitness landscapes are conducive to sub-optimal (myopic) adaptation and low average firm fitness level. We also support this relationship with findings in organisation studies and game theory and empirically show, with multiple robustness checks, that firm heterogeneity in an industry increases exit hazard. Therefore, firm heterogeneity and selection are not independent processes as often assumed in evolutionary economics. The second contribution is to uncover a hitherto unexplored indirect effect of innovation on firm survival in heterogeneous industries. We demonstrate that it is not all doom and gloom for firms in heterogeneous industries: higher levels of R&D investment in such industries enhance search for better fitness peaks and thereby mitigates the hazard-increasing effects of firm heterogeneity. Again, this mitigation effect is independent of the direct effects of R&D investment at firm or industry levels on company exit hazard.

To demonstrate why age/size heterogeneity is an additional source of exit hazard and how investment in R&D can have mitigating effect, we proceed in stages and draw on three strands of literature. We begin with evolutionary and Schumpeterian models of industry evolution (Nelson and Winter, 1982; Aghion, Akcigit, Howitt, 2014) as well as stochastic models of firm dynamics (Jovanovic, 1982; Hopenhayn, 1992; Ericson and Pakes, 1995). Then, we draw on the third strand of the literature - fitness landscape models and organisation studies - to demonstrate that industries with higher levels of firm age/size heterogeneity can approximate industry rugged fitness landscapes in evolutionary models. More rugged fitness landscapes are shown to be conducive to: (i) increased risk of remaining saddled on suboptimal fitness peaks in the presence of multiple local fitness maxima; (ii) higher costs of searching for better fitness peaks through adaptive steps of trial and error; and (iii) lower average fitness across the population (Kauffman, 1993).[2] Hence, in addition to the stochastic models of firm dynamics with passive and active learning (Jovanovic, 1982; Hopenhayn, 1992; Ericson and Pakes, 1995), we consider fitness landscape models that allow for 'cognitive leaps' towards higher fitness peaks (Gavetti and Levinthal, 2000; Felin et al., 2014) and infer that firms in industries with rugged fitness landscape can enhance searching and adaptation by investment in R&D.

The rest of the paper is organised as follows. In section 2, we review the literature summarised above with a view to derive our testable hypotheses. In section 3, we present our dataset, introduce the measures of age/size heterogeneity and discuss the estimation method. As measures of intra-industry age/size heterogeneity, we use the Theil entropy index (TI) and the coefficient of variation (CV) of firm age and employment within 3-digit SIC industries. We

[2] For rugged fitness landscape models in evolutionary biology, see Kauffman (1993 and 2016); Clune et al. (2008); De Visser and Krug (2014); and Kaznatcheev (2019). For discussion of heterogeneity effects in organisation studies, see Barnett and Hansen (1996); Barnett and Sorenson (2002); and Barnett (2008).

use a range of discrete-time hazard estimators with and without frailty (unobserved heterogeneity). Section 4 presents the results of testing two hypotheses developed in section 2. The estimated parameters indicate that firms in industries with higher age and size heterogeneity have higher rates of exit hazard, which can be mitigated through higher levels of R&D investment. These findings remain consistent across different hazard estimators, stepwise model specifications, and various firm cohorts in terms of age, size, sector, and year of entry into the industry. They also remain robust to the direct effects of age, size, R&D intensity, the entry rate in the industry, firm productivity, market concentration, and other indicators included in firm survival models.

2. Intra-industry age/size heterogeneity and learning on rugged fitness landscapes: Implications for exit hazard modelling

Malerba et al., (2016) and Capone et al., (2019) summarize a vast empirical evidence to point out persistent heterogeneity of firms by age and size across and within industries and over time, even in mature industries. An earlier insight into firm heterogeneity is suggested by Nelson and Winter (1982) - firms with bounded rationality may acquire different technological capabilities, which lead to heterogeneous behaviour and performance. Metcalfe (1998) and Foster and Hölzl (2004) emphasize the central role of firm heterogeneity as the driving force of technological and economic changes in evolutionary economics as heterogeneous companies are better in discovering and exploring new opportunities: no company variety entails stagnant industrial evolution. Malerba and Pisano (2019) summarize recent empirical research that shows persistent heterogeneity across firms by size, age, productivity, and innovativeness. These authors conclude that one of the key reasons for such heterogeneity is that technologies, processes, and products often follow trajectories with repetitive application of a fixed set of heuristics so that heterogeneous companies learn and develop capabilities differently.

Three perspectives accord firm heterogeneity a central role in the analysis of firm entry and post-entry performance: (i) Schumpeterian models of innovation and growth (Aghion and Howitt, 1992; Aghion, Akcigit, Howitt, 2014); (ii) industry evolution models with passive or active learning (Jovanovic, 1982; Hopenhayn, 1992; Ericson and Pakes, 1995); and (iii) evolutionary models that follow Nelson and Winter (1982). These theoretical models converge on three predictions, which enjoy considerable empirical support: (i) firms grow and mature as a result of past success driven by innovation and learning; (ii) firm age and size are correlated; and (iii) younger, hence, smaller firms are more likely to exit, but those that survive grow faster than the average firm in the industry (Geroski, 1995; Manjón-Antolín and Arauzo-Carod, 2008).

These predictions and the empirical support they enjoy raise two questions: the first is whether firm age and size are appropriate proxies for firm fitness, defined as the ability to survive in the face of competition and shocks. The second question, which is usually overlooked, is whether firm age/size heterogeneity can reflect an underlying fitness landscape ruggedness in the industry, with implications for selection pressure and industry evolution.

The relevance of firm size as a fitness indicator can be established from Schumpeterian models of innovation and growth (Aghion and Howitt, 1992; Aghion, Akcigit, Howitt, 2014). In these models, the firm grows and survives as a result of past success in converting innovation inputs into new product lines that increase the firm's sales and value. A similar link can be derived from both passive and active learning models of firm dynamics, where firms enter the industry with imperfect information about their true types in terms of efficiency/productivity. Efficient firms with successful investments survive and grow, others exit when payoff of exiting is higher than from remining (Jovanovic, 1982; Hopenhayn, 1992; Ericson and Pakes, 1995). Jovanovic' model assumes that firms discover their relative productivity passively before they decide whether to exit. In contrast, Ericson and Pakes (1995) assume that firms discover their productivity from their investment in research and exploration.

Therefore, the distribution of firm sizes reflects the companies' productivity draws in the passive learning models, or the idiosyncratic investment success rates in the active learning model. In both Schumpeterian models and stochastic learning models, firm size is an indicator of fitness underpinned by past success, and firm size and age are correlated. Geroski (1995) demonstrates relevant empirical evidence for this correlation. Therefore, age and size heterogeneity in the industry is an indicator of firm fitness heterogeneity, which reflects the industry's rugged fitness landscape. Such landscape, in turn, can be an additional source of exit hazard.

A fitness landscape is modelled with three parameters (Kauffman, 1993; Levinthal, 1997): the number of strategies/actions available for the firm (*N*); the number of interconnections (interdependencies) between the strategies (*K*); and the number of co-evolving links between firms/agents (*C*). These parameters determine the complexity (ruggedness) of the fitness landscape, the relevance of which for organisation studies and economics research has been acknowledged (see Krugman, 1996; Westhoff, Yarbrough, Yarbrough, 1996; Levinthal, 1997; Levinthal and Posen, 2007; Lenox, Rockart, Lewin, 2007; Khraisha, 2019; and Baumann, Schmidt, Stieglitz, 2018). Campos and Fontanari (2019) estimate that for rugged landscapes with *K* above three, the imitation strategy of another firm's actions is not optimal and agents should explore the landscape independently. The myopic exploration entails that a firm accepts a strategy change if it leads to higher fitness, however, the probability of a given path leading to monotonic improvement in fitness declines exponentially with *N* (Franke *et al*., 2011).

One common finding in this fitness landscape literature is that the search for optimal fitness or the quest for the best combination of firm strategies become increasingly harder as the ruggedness of the fitness landscape increases with firm's number of strategies (*N*) and interdependence (*K*) between strategies (Kauffman, 1993; Levinthal, 1997; Levinthal and Posen, 2007). Another common finding is that landscape ruggedness perpetuates heterogeneity among agents due to increasing difficulty of success imitation. It is assumed that firms follow myopic walk of switching on/off few strategies at a time, and such costly trial and error experimentation can trap companies with no immediate steps to higher fitness peaks under incomplete emulation of the unknown set of the best strategies and interdependencies (Rivkin, 2000; Lenox, Rockart, Lewin, 2007; 2010). Hence, evolutionary walks on rugged fitness landscapes "… can be expected to become trapped fairly quickly at local optima, rather than at the global optimum." (Macken, Hagan, Perelson, 1991, 799). Gerald et al., (2019) show that if a fraction of accessible fitness levels for a single search step is below a critical level, then the population is certain to be trapped in local optima.

Furthermore, the average fitness level is lower and the probability of a successful move to a higher fitness peaks declines as the fitness landscape becomes more rugged (Clune et al., 2008; Hartl, 2014; Gerald et al., 2019; Greenfield and Aleti, 2017). The expected number of local fitness maxima is a common measure of ruggedness, which is associated with the difficulty of adaptation on the landscape, and it increases exponentially with N, while variance of distance to the best peak is proportional to N (Neidhart, Szendro, Krug, 2014)[3]. Specifically, the probability of a successful move to a higher fitness peak declines exponentially with N (Hwang et al., 2018), while significant changes in the combination of strategies increase risk of organizational failure (Probst and Raisch, 2005). Finally, agents may not even reach local fitness optima as the ruggedness of the landscape creates path-dependency (Kaznatcheev, 2019), and agent heterogeneity could make discontinuous trajectory and evolution more likely (Andriani and Cattani, 2016; Kauffman, 2016).

[3] However, the quantitative research on the relationships between the ruggedness and "…evolutionary trajectories is still in its infancy" (Lobkovsky and Koonin, 2012, 6) and an extensive rigorous justification is quite limited in the literature. Studies using real companies' data to characterise underlying fitness landscape ruggedness (complexity) are rare as grasping this relationship is challenging. For example, Lenox, Rockart, and Lewin (2010) in a seminal study use a cross-section survey of over fourteen hundred R&D labs in the US manufacturing. Lenox with co-authors show that effect of firm interdependencies on average industry profitability is similar in magnitude to patent protection and industry growth rate. However, the interdependence was proxied by managers' subjective perceptions about complexity of their lab's processes, products, and imitation of innovations for a specific year.

On the rugged landscape the firms' independent exploration trajectories of myopic search necessary diverge leading to fitness heterogeneity, which translates into different firms' survival and growth time, hence, into heterogeneous size and age. Importantly, Lobkovsky, Wolf, and Koonin (2011) show that the measures of fitness landscape ruggedness (complexity) predict the mean path divergence on the fitness landscapes. As this path divergence translates into firms' survival time and growth heterogeneity, then the firms' heterogeneity should be associated with the fitness landscape ruggedness. The higher ruggedness of the landscape, the higher is the probability of exit by firms trapped in low local fitness optima that falls short of ensuring survival in the face of exogenous or endogenous (industry) shocks – the latter are continuously generated by companies changing their strategy combinations.

Given these theoretical findings, we propose to test the following hypothesis:

> ***H1***: *Higher level of firm age or size heterogeneity within an industry reflects a more rugged fitness landscape, where the exit hazard is higher due to higher costs of searching for optimal global fitness, lower average firm fitness level, and a higher proportion of firms stranded on local fitness peaks.*

H1 is compatible with insights from the organisation studies, where bounded rationality of companies, competency traps, and organizational inertia also decrease company survival (Hannan and Freeman, 1989; Hannan, 2005; Barnett and Hansen, 1996; Barnett and Poinikes, 2008; Levitt and March, 1988). In Barnett (2008), firms' cohort variety increases the costs of discovering and adaptation to rivals' behaviour, creates uncertainty in the firms' coevolution, and increases exit hazard. In Barnett and Hansen (1996), "Red Queen" competition increases the risk of maladaptation when firms compete with varied cohorts of rivals that have unshared co-evolutionary histories.

H1 is also compatible with predictions of game theory, where higher agents' type heterogeneity often constitutes a more complex strategy space. When a player faces unknown types of rivals with multiple levels of bounded cognitive abilities, the cost of information processing increases together with risks of choosing sub-optimal strategies (Challet and Zhang, 1998; Camerer, Ho, Chong, 2004; Kets, 2012; Strzalecki, 2014).

The analysis so far has been about how the ruggedness of the fitness landscapes, combined with bounded rationality - limited cognitive ability of firms, can be an additional source of selection pressure driven by suboptimal adaptation. However, firm search for better fitness peaks is characterised by experimental learning and environmental feedback, which can be improved (Levinthal, 2011). Gavetti (2011) has argued that learning on rugged fitness landscapes may also involve "cognitive leaps" when companies can search less myopically at larger distances. The question in this line of research is how some portions of the landscape can be discovered as a result of such leaps, especially if myopic search is less successful[4].

R&D effort is often associated with exploration in product, process, technology, and strategy spaces (Chen, 2008; Fleming and Sorenson, 2001). Companies can develop the capacity for better search and discovery of opportunities (or better fitness peaks in the NKC model) with R&D investment (Cohen and Levinthal, 1990; Bosch, Volberda, Boer, 1999; Ganco, 2017). Levinthal (1997) argues that exploration (R&D) can lead to path breaking or "long jumps" on fitness landscape. This is in line with Knudsen and Levinthal (2007) who argue that investment in exploration relaxes the cognitive constraints that causes firms to be stranded on low fitness peaks. Colombelli, Krafft, Quatraro (2013) show that variety of French manufacturing firms' patent portfolios reflect broader companies' R&D strategies, and

[4] Laboratory experiments with humans show that failures caused by increasing complexity of the environment motivates for more exploratory search (Billinger, Stieglitz, Schumacher, 2014, 103).

such variety is positively related to firms' survival due to better search of the technology landscape[5].

This resonates with active learning models where firms invest in R&D to discover their optimal market/technology niches (Ericson and Pakes, 1995). In this literature, the probability of success depends on: (i) the stochastic outcome of the R&D investment; (ii) the success of other firms in the industry; and (iii) the competitive pressure. The model predicts high mortality rates in the initial learning period, followed by longer survival for R&D-active firms that grow in size and fitness. The added value of the active learning models is that they suggest that the cognitive leaps are conditional on investment in learning and exploration, which is often associated with R&D. This matches with the finding of Lenox, Rockart, and Lewin (2007) that firms with a learning advantage are more successful in locating higher fitness peaks and their survival becomes more linked to "successful long-jumps" or "re-orientations". Given these insights, we state our second hypothesis (**H2**) as follows:

***H2**: Firms in industries with higher age/size heterogeneity can mitigate the hazard-increasing effect of heterogeneity on survival by increasing their investment in R&D.*

H2 is in line with conceptualisation of innovation as a "creative endogenous response" and "an emergent property of a complex evolving system" (Antonelli, 2009, 629) when firms face large variation and uncertainty in fitness trajectories. It is also consistent with the work indicating that environmental uncertainty increases the need for fast adaptation and innovation as reported by Covin and Slevin (1989) with respect to small firms, and by Zahra and Bogner (1999) in relation to product and service innovation in the software industry. Finally, it is also in

[5] R&D expenditures as innovation input are important for firm survival: innovating companies survive about 11% longer than those without innovation (Cefis and Marsili, 2019), and this survival premium is even larger for young and small companies (Cefis and Marsili, 2006), or when companies combine product and process innovation (Cefis and Marsili, 2012). A variety of firm key financial performance indicators have positive association with innovation (Expósito and Sanchis-Llopis, 2019). Success of start-ups also critically depend on innovation (Colombelli, Krafft, and Vivarelli, 2016). However, in industries with low technological opportunity, there is little incentive to invest in R&D (Sutton, 1998). This may partly explain the observed pattern (Figure 2) that R&D intensity is increasing in heterogeneous industries as different firms may better explore technologies.

accordance with Teece (2007), where firm performance depends on the acquisition of dynamic capabilities "to sense, seize, and reconfigure" opportunities. Certainly, dynamic capabilities are much broader than R&D efforts, but might be proxied by R&D investment.

3. Data and empirical strategy.

Our dataset is constructed by merging two UK firm-level databases: The Business Structure Database (BSD) and the Business Expenditure on Research and Development (BERD)[6]. BSD consists of the universe of UK firms registered for VAT and/or PAYE (pay-as-you-earn) purposes; and provides firm-level demographic information together with unique firm identifiers (*entref*) that allow for merging with BERD. We use the information in BSD to identify firms that exit due to bankruptcy or liquidation. The exit indicator takes the value of 1 in the year the firm exist and remain as such in the following years during which the exiting firm may remain in the register due to recording errors. We have constructed the exit year as the earliest of the death year recorded by the ONS, or the first year when the firm employment and turnover are zero for three consecutive years[7]. We also excluded firms with birth date before 1974 as firms were given the same birth year of 1973 when the business register was first introduced in 1973.

On the other hand, BERD consists of repeated annual surveys with stratified sampling of firms known to be R&D-active. The most R&D-intensive 400 firms receive a long questionnaire, with detailed questions on R&D types and sources of funding. Other firms receive a short questionnaire with questions on total, intramural and extramural R&D only. Missing data is imputed using other sources such as R&D Tax Credit returns or Annual

[6] Office for National Statistics (2019). Business Expenditure on Research and Development, 1995-2017: Secure Access [data collection]. 8th Edition. UK data Service. SN: 6690, http://doi.org/10.5255/UKDA-SN-6690-8; and Office for National Statistics (2019). Business Structure Database, 1997-2018: Secure Access [data collection]. 10th Edition. UK data Service. SN: 6697, http://doi.org/10.5255/UKDA-SN-6697-10

[7] The second criterion is used because we have identified delays in the ONS assignment of a death code in some cases even though the firm's return for employment and turnover is zero for several years.

Business Surveys. Further information on the datasets and the cleaning procedure is provided in Ugur, Trushin, Solomon (2016b).

In our dataset, the ratio of R&D to turnover is greater than one from the 96th percentile onwards. We have considered firms in the top four percent of the R&D intensity distribution as atypical and set our baseline estimation sample for firms with R&D intensity less than one. Our estimation sample consists of 35,136 firms and 158,316 observations from 1998 to 2012, of which 28,287 firms are survivors and 6,849 firms are exiters.[8] Summary statistics for the estimation sample are presented in Table OA1 in the online *Appendix* broken down by exiting firms and survivors.

Scatter plots in Figure 1 are based on data within SIC 3-digit industries and allow for inspection of the relationship between average survival times and the age/size heterogeneity in an industry. The plots in panels (a) and (b) both indicate a negative relationship between firms' age/size heterogeneity and survival time, which is equivalent to the positive association between industry heterogeneity and exit hazard rate postulated in **H1**.

[Insert Figure 1 here]

We use two measures of age and size heterogeneity as proxies for fitness heterogeneity in 265 three-digit SIC industries: the Theil entropy index and the coefficient of variation of firm age and size. In our sample, the correlation between firm age and size is 0.51 and significant, reflecting the empirical patterns discussed in section 2 (see Geroski, 1995 for a review). In what follows, we first elaborate on the advantages and drawbacks of the proposed heterogeneity measures. Then, we discuss the specification and estimation issues related to

[8] However, our results are robust to different cut-off points for R&D intensity. These results are not presented here to save space, but they are available on request.

discrete-time hazard models, the way in which we choose between estimators, and the range of sensitivity/robustness checks we conduct.

We utilise two measures or age/size heterogeneity with a set of desirable properties: the *Theil's entropy index* (*TI*) and the *coefficient of variation (CV)* of the firm age and employment sizes within 265 industries at 3-digit SIC level. The Theil entropy index for each industry/year (TI_{jt}) is calculated as:

$$TI_{jt}=\frac{1}{N_{jt}}\sum_{i=1}^{N_{jt}}\frac{L_{ijt}}{\overline{L_{jt}}}*ln\frac{L_{ijt}}{\overline{L_{jt}}} \qquad (1a)$$

Here L_{ijt} is age or employment size of the i^{th} firm in industry *j* and year *t*; $\overline{L_{jt}}$ is average age or employment in industry *j* in year *t*; and N_{jt} is the number of firms in industry *j* and year *t*. Our choice of Theil index is informed by its property of being invariant not only to unit of measurement, but also to any scale factor. The *TI* is comparable over time and between industries; and it is also additive, symmetric, and decomposable (Theil, 1972; Haughton and Khandker, 2009). Nevertheless, the *TI* is sensitive to the left end of the size distribution and may better reflect the heterogeneity among smaller firms (Haughton and Khandker, 2009).

Our second heterogeneity measure is the coefficient of variation (CV_{jt}) for firm age or employment within industry *j:*

$$CV_{jt} = \sqrt{\frac{\sum_{i=1}^{N_{jt}}(L_{ijt}-L_{jt})^2}{N_{ijt}-1}}\,\frac{1}{\overline{L_{jt}}} = S_{jt}\,\frac{1}{\overline{L_{jt}}} \qquad (1b)$$

Here S_{jt} is the standard deviation of firm age or employment in *j*-th industry, and $\frac{1}{\overline{L_{jt}}}$ is the inverse of the mean age or employment in the industry. Like *TI*, the *CV* is also invariant to multiplicative scale factors and units of measurement. The drawback here is that it is an interaction term between two variables: the standard deviation of size/age and the inverse of the mean employment or age in the industry. Therefore, we control for mean employment in the industry to avoid the risk of omitted variable bias (Sørensen and Stuart, 2000; Solanas et al., 2012). Both *TI* and *CV* are monotonically increasing with firm heterogeneity.

Scatter plots in Figure 2 shed light on a different empirical pattern in our data. Whether we measure heterogeneity with the Theil index (panel a) or the coefficient of variation (panel b), we observe a positive relationship between intra-industry age/size heterogeneity and average R&D intensity in the industry. These findings provide an empirical underpinning for H2 as it points to a higher R&D effort in more heterogeneous industries with higher risk of exit hazard. It also confirms findings in prior work, which report that firms do invest more in R&D when uncertainty is higher (Covin and Slevin, 1989; Zahra and Bogner, 1999).

[Insert Figure 2 here]

Modelling exit hazard: Main variables

We conduct hazard estimations with a view to verify if the descriptive evidence in Figures 1 and 2 is statistically significant after controlling for a wide range of firm, industry and macroeconomic factors that have been investigated in the prior literature on firm survival. To do this, we follow the general specification for the hazard rate function (Jenkins, 1995), but we use lagged, hence, predetermined or weakly exogenous covariates x_{it} to deal with simultaneity bias. The probability (*Pr*) of exit in year *t+1* conditional on observable covariates can be stated as follows:

$$Pr(y_{it+1}|x_{it,}v_i) = Pr(\beta x_{it} + \alpha M_{it} + v_i + \gamma_{t+1} + \varepsilon_{it+1}) \qquad (2)$$

Here, *i* and *t* are firm and year indices; x_{it} is a vector of observable firm-level covariates that affect firm exit with an estimated vector of β parameters; M_{it} is a vector of industry, technology (Pavitt) classes (Pavitt, 1984), and macroeconomic variables that affect firm exit with an estimated vector of α parameters; γ_{t+1} are year dummies; and ε_{it+1} is the disturbance term. The unobserved heterogeneity between firms is captured by the independently and identically distributed (i.i.d.) normal random variable $v_i|x_{it}, M_{it+1} \sim N(0,\sigma_v^2)$[9]. The strong and very common assumption in estimation of such models is that unobserved heterogeneity (v_i) and the disturbance term ε_{it+1} are independent of the firm, industry, and macroeconomic characteristics.

The variables of main interest in x_{it} include the *TI* or *CV* measures of form heterogeneity and the interactions of the latter with firm-level R&D intensity. The remaining firm-level covariates in x_{it} and also the industry, technology class and macroeconomic covariates in M_{it} are specified in accordance with the best practice in survival analysis. Definitions of all covariates and the literature that justifies their inclusion in the model are presented in Table 1.

[Insert Table 1 here]

The correlations between *TI* and *CV* is 0.44 for firm sizes and 0.37 for firm ages and both are statistically significant, but their correlation with the Herfindahl index is low and statistically insignificant, which reduces the risk of collinearity. We control for firm age and size separately in line with both theoretical and empirical work (Geroski, 1995; Klette and Kortum, 2004; Aghion et al., 2014). Although the existing literature tends to adopt a linear specification for the effects of innovation on survival, Ugur, Trushin, Solomon (2016a) demonstrate that a quadratic specification could be more plausible both theoretically and empirically.

[9] How important is the assumption that unobserved heterogeneity v_i is i.i.d. Normal? Nicoletti and Rondinelli (2010) have evaluated biases in estimated parameters of the discrete time hazard models caused by omitting or misspecifying the unobserved heterogeneity distribution using Monte-Carlo simulations. Their results demonstrate that misspecification of the unobserved heterogeneity is unlikely to lead to a significant bias.

We also control for labour productivity and firm growth rate relative to industry growth, which are reported as significant determinants of firm survival (Audretsch, 1991, 1995; Hopenhayn, 1992; Ericson and Pakes, 1995; Mata, Portugal, Guimaraes, 1995; Cefis and Marsili, 2005; and Ugur, Trushin, Solomon, 2016a). The other set of firm-level characteristics includes number of plants, whether the firm is engaged in civil R&D only, and domestic versus foreign ownership reported as significant by Audretsch (1991), Audretsch and Mahmood (1995), Mata, Portugal, Guimaraes, (1995); and Fernandes and Paunov (2015).

Of the industry-level covariates, Baldwin and Rafiquzzaman (1995) report that higher entry rates tend to reduce firm survival. Positive association between entry and exit rates at the industry level has been reported by Dunne, Roberts, Samuelson (1988). However, if the landscape complexity is high, the negative effects of firm entry into an industry tend to be lower in simulations of the NK model (Wu et al., 2019). The effect of market concentration also varies, but it tends to be insignificant (see Baldwin and Rafiquzzaman, 1995; McCloughan and Stone, 1998; Ugur, Trushin, Solomon, 2016a).

We control for average number of firm employees at 3-digit SIC industry level in order to address the risk of omitted variable bias, and include the industry median R&D intensity to verify whether higher level of creative destruction affects firm mortality (Aghion *et al*., 2014; Fernandes and Paunov, 2015; Ugur, Trushin, Solomon, 2016a). Finally, we check if technology classes matter using the Pavitt (1984) industrial technology taxonomy.[10] The final set of covariates relates to macro-level indicators such as onsets of the financial crisis, real effective exchange rate of British pound, and GDP growth. Whilst currency appreciation may affect mortality through decline in international cost competitiveness, the recent financial crisis

[10] Pavitt technology classes are revised slightly by Bogliacino and Pianta (2010). Pavitt1 consists of firms in science-based industries such as chemicals, office machinery, precision, medical and optical instruments industries, ICT. Pavitt2 includes specialized suppliers of technology - mechanical engineering industries, manufacturers of electrical machinery, equipment, etc. Pavitt3 includes scale-intensive industries such as pulp and paper, transport vehicles, mineral oil refining industries. Pavitt5 consists of unclassified industries.

dummy accounts for changes in the business and credit environment. Finally, GDP growth captures the effect of business cycle on firm survival (Goudie and Meeks, 1991; Bhattacharjee *et al*., 2009; Ugur, Trushin, Solomon, 2016a).

Estimation methodology

Our estimation methodology follows Wooldridge (2010) on grouped duration data, where firm exit time is known within one year. The discrete-time hazard rate h_{it} that firm *i* exits in T_e years conditional on survival for T_{e-1} years can be expressed as conditional probability of firm survival for T_i years as follows:

$$h_{it} = \frac{\Pr(T_{e-1} < T_i \leq T_e)}{\Pr(T_i > T_{e-1})} \tag{3}$$

Bauer and Agarwal (2014, 432) provide evidence that hazard models are "superior to the alternatives" in the context of estimating bankruptcy hazards. The parameters are estimated by maximizing the logarithm of the likelihood function. Whereas the Logit specification assumes a logistic distribution for the hazard, the Probit assumes a standard Normal distribution. Given the panel structure of the data, we choose random effect estimations as it helps to correct for omitted variable bias (Fernandes and Paunov, 2015), whereas fixed-effect estimations often lead to large biases in all estimated parameters with relatively small number of periods in the dataset due to the incidental parameters problem (Lancaster, 2000; Wooldridge, 2010; Bester and Hansen, 2009). The dependent variable is an indicator taking the value of one if the firm exits in year T_e, and zero otherwise. To partially eliminate competing causal attributions, we use one-year forward firm exit as our dependent variable (formula 2).

The panel random effect estimator controls for unobserved firm heterogeneity[11]. Geroski, Mata, Portugal (2010) emphasize the importance of such control. Wooldridge (2010)

[11] For the random effects model the maximum log-likelihood estimations are based on Gauss-Hermite quadrature approximation (see Naylor and Smith, 1982) with a corresponding probability distribution hazard function *Pr(z)*. To check for robustness, we use both non-proportional hazard functions (Logit and Probit) and

demonstrates that a $\sqrt{N}$ consistent estimator in this case, the population-averaged parameters, can be obtained by maximization of the log-likelihood function $log\ L$:

$$log\ L = \sum_{i=1}^{N} \sum_{t=1}^{Te} \{y_{it} log\ h_{it} + (1- y_{it}) log\ (1 - h_{it})\} \quad (4)$$

Typical distribution specification for the random-effects estimators are given by the standard Normal Φ cumulative density functions by Wooldridge (2010):

$$Probit: Pr(z) = \Phi(z) = \Phi(x_{it}\beta + M_{it+1}\alpha + v_i + \gamma_{t+1} + \varepsilon_{it+1}) \quad (5a)$$

$$Logit: Pr(z) = 1/(1 + exp(-z)) \quad (5b)$$

In the complementary log-log random-effects estimator, the conditional probability function is given by *Clog-log*: $Pr(z) = 1 - exp\ (-\,exp(z))$ (5c)

We use likelihood ratio test to check if the panel random effects estimators deliver similar results with the pooled estimators - if the panel level variance is insignificant and the ratio $\rho = \frac{\sigma_v{}^2}{\sigma_\varepsilon{}^2 + \sigma_v{}^2}$ is different from zero by sampling chance[12]. We also use robust standard errors of the estimated parameters, which provide consistency when the disturbances are not correlated across firms.

Importantly, in nonlinear models the interaction effect is not equivalent to the marginal effect - the sign of the estimated parameter for the interaction term between R&D intensity and the firm heterogeneity indicators within an industry can be misleading (Norton, Wang, Ai, 2004). We numerically estimate the interaction effects by using *margins* (Williams, 2012) and *inteff* (Norton, Wang, Ai, 2004) procedures in Stata based on delta approximation method applied to Probit model, which is selected by the AIC and BIC information criteria[13]. In this model, the interaction effect for the conditional mean of the indicator variable *y* is:

---

the proportional specification through Complementary log-log (Clog-log). Although Jenkins (1995) notes that both estimators tend to converge when hazard rates are small, it is appropriate to use both types of estimators as the hazard functions are not known *ex ante*.

[12] Stata reports panel level variance $\ln(\sigma_v{}^2)$ in form of lnsig2u_const.

[13] Akaike (AIC) and Bayesian (BIC) information criteria help to choose between non-nested models by asymptotically minimizing information loss – the lowest criterion favours more parsimonious model. The criteria are estimated as -2logL + *c*p, where L is the likelihood function, *k* is the number of parameters in the

$$E[y|x_1, x_2, X] = \Phi(\alpha_1 x_{1t} + \alpha_2 x_{2t} + \alpha_{12} x_{1t} x_{2t} + x_{jt}\alpha_j) = \Phi(z) \quad (6)$$

According to Norton, Wang, Ai (2004), the full marginal effect of the interaction term on the conditional mean survival is:

$$\frac{\partial^2 \Phi(z)}{\partial x_1 x_2} = [\alpha_{12} - (\alpha_1 + \alpha_{12} x_2)(\alpha_2 + \alpha_{12} x_1) z]\, \Phi'(z) \quad (7)$$

Hence, the marginal effect of the interaction term depends on specific levels of all covariates. We also report graphical representations of the estimated interaction effects following Greene's (2010) recommendation for nonlinear models.

4. Results and discussion

We report results form a wide range of discrete-time hazard models: pooled Probit, Logit, Complementary log-log (Clog-log); and their panel random effects versions. We present the preferred estimation results in the main text; and the additional robustness checks in the online *Appendix*. The preferred estimators are determined by AIC/BIC values, which point in favour of Probit estimator. Table 2 reports pooled and random-effects Probit estimations for both measures of size heterogeneity: the Theil entropy index (columns 1 and 2) and the coefficient of variation (columns 3 and 4) of firm employment. We have also conducted the likelihood ratio (LR) test to check whether the panel random-effects estimators are preferable to pooled estimators - the test favours the random effects estimator, which we use to obtain non-linear interaction marginal effects (Table 3) and conditional effects depending on various levels of R&D intensity (Table 4).[14]

model, and the coefficient *c* is 2 for AIC and logarithm of the number of observations for BIC (Aho, Derryberry, Peterson, 2014).

[14] Estimation results for age heterogeneity and exit hazards are presented in the online *Appendix*. The estimated parameters are fully consistent with those based on the size heterogeneity. The LR tests and the industry clustered standard errors are not reported to save space as they confirm the results in the main text and can be provided on request.

[Insert Table 2 here]

Post-estimations for pooled Probit indicate that: (i) the model fits the data well as the Pearson χ^2 does not reject the null hypothesis of good fit; (ii) the overall rate of correct classification is high, at 95.66% in the estimation based on the Theil index and 95.67% for the coefficient of variation; and (iii) the model has good power to discriminate between exiting and surviving firm as the area under the ROC curve is 0.69 and 0.68, respectively. There is sign and significance consistency across six estimators and two heterogeneity measures. The consistency is evident with respect to covariates of main interest (both heterogeneity measures, their interactions with R&D intensity, and the latter's linear and quadratic terms), and the wide range of controls in the firm survival literature. Furthermore, estimated parameters (reported in the online Appendix) are robust across step-wise estimations, all age and size heterogeneity measures, various firm cohorts and sectors.

The estimated parameters for age/size heterogeneity are significant and provide strong support for H1. This is observed after controlling for logarithm of firm size and its square, and also for the mean employment in the industry. Hence, the intra-industry size heterogeneity is a reasonable proxy for the landscape ruggedness, which is an exit hazard in its own right. The estimated effects are robust after controlling for firm size and mean size in the industry in order to minimize the risk of bias due to the left-tail sensitivity of the Theil index and the omitted variable bias that may be caused by the coefficient of variation (Stirling, 2010; Solanas *et al.*, 2012). The parameter estimates for the Theil entropy index are larger than those of the coefficient of variation. This, however, is to be expected because the mean of the coefficient of variation in the sample is 5.0 compared to a mean Theil index of 1.9, hence, the parameter estimates are consistent in terms of sign and magnitudes.

The negative and statistically significant parameter estimates for the interaction between the heterogeneities and R&D intensity lend support to H2, which posits that higher levels of R&D investment counterbalance the adverse effect of the landscape complexity on firm survival. We take account of the non-linear nature of the hazard estimators and obtain numerical estimates of the parameter using *margins* and *inteff* procedures in *Stata* (Norton, Wang, Ai, 2004; Williams, 2012). We also provide graphical representations of the estimated interaction effects following Greene's (2010) recommendation for nonlinear models. Drawing on Greene (2010), we present in Figure 3 the sign and significance of the interaction effects from Probit model. The horizontal axis indicates the predicted probability of exit, whereas the vertical axis indicates the associated Z-statistics. The horizontal lines above and below zero demarcates the Z-values that corresponds to statistical significance at 5%. It can be seen that the parameter estimates for the interaction term are associated with negative Z-statistics, which confirm the significant negative interaction effects as the Z-statistics are almost always below the demarcation line for significance. These estimations support H2 that the R&D intensity ameliorates the adverse effect of the landscape ruggedness on firm survival.

[Insert Figure 3 here]

Table 3 provides further evidence to support H2 by using the delta method and routines presented by Norton, Wang, Ai, (2004) and Williams (2012). The estimated marginal interaction effects are always negative and significant at mean values of all covariates. When the full range of the covariate values are taken into account, the marginal effects are predominantly negative and significant. Hence, we can safely conclude that R&D intensity diminishes effects of the landscape complexity (as proxied by industry firm size/age heterogeneity) on exit hazard. These findings tie in with the descriptive evidence presented by Figure 1 and 2.

[Insert Table 3 here]

The estimated parameters are strongly consistent across various sub-samples (see the online Appendix), including: (i) alternative pooled and panel hazard estimators; (ii) different cut-off points for R&D intensity; (iii) step-wise regressions; (iv) samples that avoid left truncation by investigating firms born in or after 2000 or 2003; (v) samples that exclude firms in the financial or defence industries. Vivarelli (2013) points to heterogeneous capabilities of new entrepreneurs as some companies might be established to escape unemployment by low business capacity founders, which could make more mistakes and fail earlier. Bartelsman, Scarpetta, and Schivardi (2005) found out that about 20-40% of entering firms fail within two years in ten OECD countries. To exclude the possibility that we observe this firm heterogeneity effect for young and small firms only, we have estimated the hazard models for firms above the sample average age and size (see online Attachment Table OA11) - the results are consistent with the full-sample estimations. This finding indicates that the adverse effect of age/size heterogeneity on survival in the full sample is not driven by small or young firms.

The estimated parameters for other covariates are in line with the findings in the literature (Table 1). First, we confirm the diminishing scale effect in the relationship between R&D intensity and survival reported earlier in Ugur, Trushin, Solomon, (2016a) for hazard models specified in continuous time. The U-shaped relationship between R&D intensity and exit hazard can be due to either increased riskiness of R&D investments at higher levels of R&D intensity (Ericson and Pakes, 1995; Czarnitzki and Toole, 2013) or firms' failure to diversify their revenue at the same pace as their investment in R&D (Fernandes and Paunov, 2015; Ugur, Trushin, Solomon, 2016a).

The estimated parameters also confirm that new entrants have shorter survival time (Klette and Kortum, 2004; Aghion *et al.*, 2014; Cefis and Marsili, 2005; Evans, 1987). The U-shaped relationship between firm's size and exit hazard is in line with the non-monotonic evidence on size distribution and survival among Portuguese firms (Cabral and Mata, 2003), which suggests that a firm size beyond an efficient scale may be a hazard factor.

Following Audretsch and Mahmood (1995) and Fernandes and Paunov (2015), we report that multi-plant firms are less likely to exit as they are better able to diversify risk. We also report that real turnover per employee and growth rates relative to median growth in the industry are associated with lower exit hazard and this confirms earlier findings of Doms, Dunne, Roberts (1995), Mata, Portugal, Guimaraes (1995), and Griliches and Regev (1995) among others.

Of the industry-level covariates, we find that the relationship between exit hazard and market concentration is insignificant. This is in line with some prior studies, which offer the following explanations: (i) market concentration may be less important than market niches in determining monopoly rents (Geroski, 1995); (ii) entry costs in concentrated industries depend on actions of hypercompetitive and less predictable firms, but not on the number of companies in an industry (D'Aveni, Dagnino, Smith, 2010); and (iii) industries with similar concentration ratios often show significant heterogeneity in the overall firm-size distribution (Carroll, 1985, 1264). Perhaps, the market concentration index is not significant in comparison to the landscape complexity indicators.

Four Pavitt classes are also significant, which is in line with empirical findings of Agarwal and Audretsch (2001) and also Cefis and Marsili (2005) that the nature of the technology in an industry impacts survival. The average R&D intensity in the industry, as a proxy for the creative destruction caused by the industry-wide level of innovation, increases the hazard in line with the Schumpeterian innovation models (Aghion, Akcigit, Howitt, 2014; Ugur, Trushin, Solomon, 2016a) and previous empirical findings (Fernandes and Paunov, 2015).

With respect to macroeconomic variables, we report that the real currency appreciation, which reduces competitiveness against foreign firms, and the onset of the recent financial crisis increase exit hazard; whereas the GDP growth rate has a negative relationship with exit hazard as the domestic demand expands. These findings are in line with those reported in firm survival studies that control for macroeconomic variables (Bhattacharjee et al., 2009; Goudie and Meeks, 1991; Ugur, Trushin, Solomon, 2016a).

The final set of evidence we present here has significant implications for organisational strategy because it sheds light on the levels of R&D intensity required to counter-balance and eventually reverse the hazard-increasing effect of the landscape complexity as proxied by the firm heterogeneity. Table 4 presents the impacts of the heterogeneity on exit hazard conditional on different levels of firm R&D intensity.

[Insert Table 4 here]

Fixing the covariates at their sample means, we varied the level of firm R&D intensity from the bottom 5th to the top 95th percentile. The results indicate that the marginal hazard effects of the landscape complexity decline as R&D intensity increases, and the effects become insignificant between the 70th and 75th percentiles of the R&D intensity and eventually reverse at the top R&D intensity[15]. In other words, firms with the sample averages characteristics find that the impact of the firm heterogeneity on exit hazard declines to zero as the firms' R&D

[15] In the case of age heterogeneity, the adverse effect is diluted significantly as R&D intensity increases, but it is never neutralised or reversed due to lower magnitude of firms' age heterogeneity, which could be less informative measure of the landscape complexity.

intensity (the ratio of R&D expenditures to turnover) achieves 70th percentile, which is about 9–11% of firm turnover.

5. Conclusion

Drawing on the fitness landscape literature and the firm dynamic models we have addressed two novel research questions: (i) are within-industry firm age and size heterogeneities associated with higher exit hazards as such heterogeneities indicate possibilities for myopic adaptation on the rugged fitness landscapes?; and (ii) can firm investment in R&D mitigate this new form of exit hazard in heterogeneous industries?

We have provided theoretical insights from evolutionary economics, biology, organisation studies, and industrial organisation as to why observed firm heterogeneity, which reveals underlying fitness landscape complexity/ruggedness, can be an additional source of selection and churning pressure due to companies' myopic adaptation. The literature review implies that firm heterogeneity is persistent not only because of continual flux of firm entry, exit, or technological differences, but also due to some firms stranded at local fitness peaks that are lower than the global maximum. We have also provided a wealth of empirical evidence supporting our hypotheses, with the evidence remaining consistent across different hazard estimators, stepwise model specifications, and different firm cohorts in terms of age, size, sector, and year of entry into the industry.

Our analysis and findings offer three contributions to the exiting knowledge on firm dynamics and industry evolution. The first is the conceptual justification and empirical testing of new measure (proxy) of the fitness landscape ruggedness with intra-industry variation in firm age and size. This proxy provides an opportunity to empirically test predictions of the NKC model with company data, beyond hypothetical numerical simulations. The second is that firm heterogeneity is not just a passive precondition for subsequent selection process as often assumed in evolutionary economics. The firms age/size heterogeneity in the industry is both an indicator of potentially sub-optimal selection and an additional source of selection pressure. As such, heterogeneity indicates a rugged (complex) fitness landscape where the average fitness level may be lower due to co-evolutionary firm cohorts stranded at low fitness peaks in their path-dependent myopic search for the globally maximum fitness level. One implication that follows from our finding is that the generic strategy of diversification proposed by Porter (1980) has unintended effects of increasing industry heterogeneity, which makes the fitness landscapes more complex and this increases exit hazard. The second implication is that R&D improves survival in heterogeneous industries by facilitating learning on rugged fitness landscapes. This is why it is not surprising to observe that firms in industries with higher age/size heterogeneity do invest more in R&D. Our findings suggest that it is good practice to control for age/size heterogeneity and for the interactions of the latter with R&D intensity on both theoretical and empirical grounds.

References

Agarwal, R., and D.B. Audretsch. 2001. “Does Entry Size Matter? The Impact of the Life Cycle and Technology on Firm Survival”. *The Journal of Industrial Economics* 49: 21-43.

Agarwal, R., and M.Gort. 2002. “Firm and Product Life Cycles and Firm Survival”. *The American Economic Review* 92(2): 184-190.

Aghion, P., and P. Howitt. 1992. A Model of Growth Through Creative Destruction. *Econometrica: Journal of the Econometric Society*, 60(2), 323-351.

Aghion, P., U. Akcigit, and P. Howitt. 2014. “What Do We Learn from Schumpeterian Growth Theory?” In *Handbook of Economic Growth* (Vol. 2), 515-563, Elsevier.

Aho, K., D. Derryberry, and T. Peterson. 2014. “Model Selection for Ecologists: The Worldviews of AIC and BIC”. Ecology, 95 (3): 631–636.

Andriani, P., and G. Cattani. 2016. “Exaptation as Source of Creativity, Innovation, and Heterogeneity: Introduction to the Special Section”. *Industrial and Corporate Change 25*(1): 115-131.

Antonelli, C. 2009. “The Economics of Innovation: From the Classical Legacies to the Economics of Complexity”. *Economics of Innovation and New Technology* 18(7): 611-646.

Audretsch, D. 1991. “New-Firm Survival and the Technological Regime”. *Review of Economics and Statistics* 60(3): 441-450.

Audretsch D.B. 1995. “Innovation, Growth and Survival”. *International Journal of Industrial Organization* 13(4): 441-457.

Audretsch D.B, P. Houweling, and A.R. Thurik. 2004. “Industry Evolution: Heterogeneity, Selection and the Role of Learning”. *International Small Business Journal* 22(4): 331-348.

Audretsch D.B., T. Mahmood. 1995. “New Firm Survival: New Results Using a Hazard Function”. *The Review of Economics and Statistics* 77(1): 97-103.

Baldwin J., and M. Rafiquzzaman. 1995. “Selection Versus Evolutionary Adaptation: Learning and Post-entry Performance”. *International Journal of Industrial Organization* 13: 501-522.

Barnett, W. 2008. *The Red Queen Among Organizations: How Competitiveness Evolves*. Princeton: Princeton University Press.

Barnett, W., and M. Hansen. 1996. “The Red Queen in Organizational Evolution”. *Strategic Management Journal* 17: 137-157.

Barnett, W., and O. Sorenson. 2002. “The Red Queen in Organizational Creation and Development”. *Industrial and Corporate Change* 12(4): 673-695.

Barnett, W., and G. Pontikes. 2008. “The Red Queen, Success Bias, and Organizational Inertia”. *Management Science* 54(7): 1237-1251.

Bartelsman, E., S. Scarpetta, and F. Schivardi. 2005. “Comparative Analysis of Firm Demographics and Survival: Evidence from Micro-level Sources in OECD Countries”. *Industrial and Corporate Change,* 14 (3): 365–91.

Bauer J, and V. Agarwal. 2014. “Are Hazard Models Superior to Traditional Bankruptcy Prediction Approaches? A Comprehensive Test”. *Journal of Banking and Finance* 40: 432-442.

Baumann O, J. Schmidt, N. Stieglitz. 2018. “Effective Search on Rugged Performance Landscapes: A Review and Outlook”. *Journal of Management* 45(1): 285–318.

Bester, A., and C. Hansen. 2009. “A Penalty Function Approach to Bias Reduction in Nonlinear Panel Models with Fixed Effects”. *Journal of Business & Economic Statistics* 27(2): 131-48.

Bhattacharjee, A., C. Higson, S. Holly, and P. Kattuman. 2009. “Macroeconomic Instability and Business Exit: Determinants of Failures and Acquisitions of UK Firms”. *Economica* 76(301): 108-131.

Billinger, S., N. Stieglitz, T.R. Schumacher. 2013. “Search on Rugged Landscapes: An Experimental Study”. *Organization Science 25*(1): 93-108.

Bogliacino, F., and M. Pianta. 2010. “Innovation and Employment: A Reinvestigation Using Revised Pavitt Classes”. *Research Policy* 39(6): 799-809.

Bosch, F., H. Volberda, M. Boer. 1999. “Coevolution of Firm Absorptive Capacity and Knowledge Environment: Organizational Forms and Combinative Capabilities”. *Organization Science* 10(5): 551-68.

Cabral, L., and J. Mata. 2003. “On the Evolution of the Firm Size Distribution: Facts and Theory”. *American Economic Review* 93(4): 1075-1090.

Camerer, C., T.-H. Ho, J.-K. Chong. 2004. “A Cognitive Hierarchy Model of Games”. *Quarterly Journal of Economics* 119(3): 861-898.

Campos, P.R., and J.F. Fontanari. 2019. “Predictability of Imitative Learning Trajectories”. *Journal of Statistical Mechanics: Theory and Experiment* 2019(1): 013501.

Capone, G., F. Malerba, R. Nelson, L. Orsenigo, S. Winter. 2019. “History Friendly Models: Retrospective and Future Perspectives”. *Eurasian Business Review* 9: 1-23.

Carroll, G.R. 1985. “Concentration and Specialization: Dynamics of Niche Width in Populations of Organizations”. *American Journal of Sociology* 90: 1262-283.

Castellacci, F. 2011. "Theoretical Models of Heterogeneity, Growth and Competitiveness: Insights from the Mainstream and Evolutionary Economics Paradigms." *International Handbook on the Economics of Integration: Competition, Spatial Location of Economic Activity and Financial Issues* 2: 90.

Cefis, E., and O.Marsili. 2005. “A Matter of Life and Death: Innovation and Firm Survival”. *Industrial and Corporate Change* 14(6): 1167-1192.

Cefis, E., and O.Marsili. 2006. “Survivor: The role of innovation in firms’ survival”. *Research Policy* 35(5): 626-641.

Cefis, E., and O.Marsili. 2012. “Going, Going, Gone. Exit Forms and the Innovative Capabilities of Firms”. *Research Policy* 41(5): 795-807.

Cefis, E., and O.Marsili. 2019. “Good Times, Bad Times: Innovation and Survival Over the Business Cycle”. *Industrial and Corporate Change* 28(3): 565-587.

Challet, D., and Y.-C.Zhang. 1998. “On the Minority Game: Analytical and Numerical Studies”. *Physica A* 256: 514-532.

Chen, W.R., 2008. “Determinants of Firms' Backward-and Forward-looking R&D Search Behavior”. *Organization Science 19*(4): 609-622.

Clune, J., D. Misevic, C. Ofria, R. Lenski, S. Elena, and R. Sanjuán. 2008. “Natural Selection Fails to Optimize Mutation Rates for Long-term Adaptation on Rugged Fitness Landscapes”. *PLoS Computational Biology 4*(9): 1-8.

Coad, A., J.R. Holm, J. Krafft, and F. Quatraro. 2018. “Firm Age and Performance”. *Journal of Evolutionary Economics 28*(1): 1-11.

Cohen, W., and D. Levinthal. 1990. “Absorptive Capacity: A New Perspective on Learning and Innovation”. *Administrative Science Quarterly* 35(1): 128-152.

Colombelli, A., J. Krafft, and F. Quatraro. 2013. "Properties of Knowledge Base and Firm Survival: Evidence from a Sample of French Manufacturing Firms". *Technological Forecasting and Social Change 80*(8): 1469-1483.

Colombelli, A., J. Krafft, M. Vivarelli. 2016. "To Be Born Is Not Enough: The Key Role of Innovative Startups". *Small Business Economics* 47: 277-291.

Covin, J., and D.Slevin. 1989. "Strategic Management of Small Firms in Hostile and Benign Environments". *Strategic Management Journal* 10: 75-87.

Czarnitzki, D. and A.A. Toole. 2013. "The R&D Investment–uncertainty Relationship: Do Strategic Rivalry and Firm Size Matter?" *Managerial and Decision Economics* 34(1): 15-28.

D'Aveni, R., G. Dagnino, and K. Smith. 2010. "The Age of Temporary Advantage". *Strategic Management Journa*l 31: 1371-85.

De Visser, J., and J. Krug. 2014. "Empirical Fitness Landscapes and the Predictability of Evolution". *Nature Reviews Genetics 15*(7): 480-490.

Disney, R., J. Haskel, Y. Heden. 2003. "Entry, Exit and Establishment Survival in UK Manufacturing". *Journal of Industrial Economics* 51(1): 91-112.

Doms M., T. Dunne, and M.J. Roberts. 1995. "The Role of Technology Use in the Survival and Growth of Manufacturing Plants". *International Journal of Industrial Organization* 13: 523-542.

Dosi, G., F. Malerba, O. Marsili, L. Orsenigo. 1997. "Industrial Structures and Dynamics: Evidence, Interpretations and Puzzles". *Industrial and Corporate Change* 6(1): 3-25.

Dunne, T., M.J. Roberts, and L. Samuelson. 1988. "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries". *RAND Journal of Economics* 19(4): 495-515.

Ericson, R., and A. Pakes. 1995. "Markov-perfect Industry Dynamics: A Framework for Empirical Work". *Review of Economic Studies* 62(1): 53-82.

Evans, D.S. 1987. "The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries". *The Journal of Industrial Economics* 35: 567-581.

Expósito, A., and J. Sanchis-Llopis. 2019. "The Relationship Between Types of Innovation and SMEs' Performance: A Multi-dimensional Empirical Assessment". *Eurasian Business Review* 9(2): 115-135.

Felin, T., S. Kauffman, R. Koppl, and G. Longo. 2014. "Economic Opportunity and Evolution: Beyond Landscapes and Bounded Rationality". *Strategic Entrepreneurship Journal* 8(4): 269-282.

Fernandes, A., and C. Paunov. 2015. "The Risks of Innovation: Are Innovating Firms Less Likely to Die?" *The Review of Economics and Statistics* 97(3): 638-53.

Fleming, L., and O. Sorenson. 2001. "Technology as a Complex Adaptive System: Evidence from Patent Data". *Research Policy* 30(7): 1019-1039.

Foster, J. and W. Hölzl, W. 2004. “Introduction and overview”. In *Applied Evolutionary Economics and Complex Systems*, Foster, J. and W. Hölzl, ed., 1-14, London: Edward Elgar Publishing.

Franke, J., A. Klözer, J. de Visser, and J. Krug. 2011. “Evolutionary Accessibility of Mutational Pathways”. *PLoS Computational Biology 7*(8): e1002134.

Ganco, M. 2017. “NK Model as a Representation of Innovative Search”. *Research Policy* 46(10): 1783-1800.

Gavetti, G. 2011. “The New Psychology of Strategic Leadership”. *Harvard Business Review* 89(7-8): 118-25.

Gavetti, G., and D. Levinthal. 2000. “Looking Forward and Looking Backward: Cognitive and Experiential Search”. *Administrative Science Quarterly* 45(1): 113-137.

Gerald, N., D. Dutta, R.G. Brajesh, and S .Saini. 2019. “Mathematical Modeling of Movement on Fitness Landscapes”. *BMC Systems Biology* 13(1): 25.

Geroski, P.A. 1995. “What Do We Know About Entry?” *International Journal of Industrial Organization* 13(4): 421-440.

Geroski, P., J. Mata, and P. Portugal. 2010. “Founding Conditions and the Survival of New Firms”. *Strategic Management Journal* 31: 510-529.

Greene, W. 2010. “Testing Hypotheses about Interaction Terms in Nonlinear Models”. *Economic Letters* 107: 291-296.

Greenfield, G., and A. Aleti. 2017. “On Pathological Fitness Landscapes for Constrained Combinatorial Optimization”. *Federated Conference on Computer Science and Information Systems* 12: 81-86.

Griliches, Z., and H. Regev. 1995. “Firm Productivity in Israeli Industry 1979–1988”. *Journal of Econometrics* 65(1): 175-203.

Goudie, A.W., and G. Meeks. 1991. “The Exchange Rate and Firm Failure in a Macro-micro Model of the UK Firm Sector”. *Economic Journal* 101: 444-457.

Hannan, M.T., and J. Freeman. 1989. *Organizational Ecology*. Cambridge, MA: Harvard University Press.

Hannan, M. 2005. “Ecologies of Organizations: Heterogeneity and Identity”. *Journal of Economic Perspectives* 19(10): 51-70.

Hartl, D.L. 2014. “What Can We Learn from Fitness Landscapes?” *Current Opinion in Microbiology 21:* 51-57.

Haughton, J., and S. Khandker. 2009. *Handbook on Poverty and Inequality*. Washington, D.C.: The World Bank.

Hofer, C. 1975. “Toward a Contingency Theory of Business Strategy”. *Academy Management Journal* 18(4): 784-810.

Hopenhayn, H. 1992. “Entry, Exit, and Firm Dynamics in Long Run Equilibrium”. *Econometrica* 60(5): 1127–1150.

Hwang, S., B. Schmiegelt, L. Ferretti, and J. Krug. 2018. “Universality Classes of Interaction Structures for NK Fitness Landscapes”. *Journal of Statistical Physics 172*(1): 226-278.

Jenkins, S. 1995. “Easy Ways to Estimate Discrete Time Duration Models”. *Oxford Bulletin of Economics and Statistics* 57: 129-138.

Jovanovic, B. 1982. “Selection and the Evolution of Industry”. *Econometrica* 50(3): 649-670.

Kauffman, S. 1993. *The Origins of Order*. Oxford: Oxford University Press.

Kauffman, S. 2016. *A Creative Universe: The Social World*. Oxford: Oxford University Press.

Kaznatcheev, A., 2019. “Computational Complexity as an Ultimate Constraint on Evolution”. *Genetics 212*(1): 245-265.

Kets, W., 2012. “Learning with Fixed Rules: The Minority Game”. *Journal of Economic Surveys* 26(5): 865-878.

Klette, T., and S. Kortum. 2004. “Innovating Firms and Aggregate Innovation”. *Journal of Political Economy 112*(5): 986-1018.

Knudsen, T., and D.A. Levinthal. 2007. “Two Faces of Search: Alternative Generation and Alternative Evaluation”. *Organization Science 18*(1): 39-54.

Khraisha, T. 2019. “Complex Economic Problems and Fitness Landscapes: Assessment and Methodological Perspectives”. *Structural Change and Economic Dynamics,* in press.

Krugman, P. 1996. *The Self-Organizing Economy*. Boston: Blackwell Publishers.

Lancaster T. 2000. “The Incidental Parameter Problem Since 1948”. *Journal of Econometrics* 95(2): 391-413.

Lenox, M.J., S.F. Rockart, and A.Y. Lewin. 2007. “Interdependency, competition, and industry dynamics”. *Management Science* 53(4): 599-615.

Lenox, M.J., S.F. Rockart, A.Y. Lewin. 2010. “Does Interdependency Affect Firm and Industry Profitability? An Empirical Test”. *Strategic Management Journal* 31(2): 121-139.

Levinthal, D. 1997. “Adaptation on Rugged Landscapes”. *Management Science 43*(7): 934-950.

Levinthal, D. 2011. “A Behavioral Approach to Strategy - What's the Alternative?”. *Strategic Management Journal* 32(13): 1517-1523.

Levinthal, D., and H.E. Posen. 2007. “Myopia of Selection: Does Organizational Adaptation Limit the Efficacy of Population Selection?” *Administrative Science Quarterly 52*(4): 586-620.

Levitt, B., and J.G. March. 1988. “Organizational Learning”. *Annual Review of Sociology 14*(1): 319-338.

Lobkovsky, A.E. and Koonin, E.V., 2012. Replaying the tape of life: quantification of the predictability of evolution. *Frontiers in Genetics*, 3, Article 246.

Lobkovsky, A.E., Y.I. Wolf, E.V. Koonin. 2011. "Predictability of Evolutionary Trajectories in Fitness Landscapes". *PLoS Computational Biology* 7(12): e1002302.

Luttmer, E., 2007. "Selection, Growth, and the Size Distribution of Firms". *Quarterly Journal of Economics* 123(3): 1103-1144.

Macken, C.A., P.S. Hagan, and A.S. Perelson. 1991. "Evolutionary Walks on Rugged Landscapes". *SIAM Journal on Applied Mathematics 51*(3): 799-827.

Malerba F., R. Nelson, L. Orsenigo, S. Winter. 2016. *Innovation and the evolution of industries-history-friendly models*. Cambridge: Cambridge University Press.

Malerba, F. and G. Pisano. 2019. "Innovation, Competition and Sectoral Evolution: An Introduction to the Special Section on Industrial Dynamics". *Industrial and Corporate Change* 28(3): 503-510.

Manjón-Antolín, M. and J. Arauzo-Carod. 2008. "Firm Survival: Methods and Evidence". *Empirica 35*(1): 1-24.

Mata, J., P. Portugal, and P. Guimaraes. 1995. "The Survival of New Plants: Start-up Conditions and Post-entry Evolution". *International Journal of Industrial Organization* 13: 459-481.

Mata, J., and P. Portugal. 2002. "The Survival of New Domestic and Foreign Owned Firms". *Strategic Management Journal* 23: 323-343.

McCloughan, P., and I. Stone. 1998. "Life Duration of Foreign Multinational Subsidiaries: Evidence from UK Northern Manufacturing Industry 1970-93". *International Journal of Industrial Organization* 16(6): 719-747.

Metcalfe, J.S. 1998. *Evolutionary economics and creative destruction*. London: Psychology Press.

Naylor, J., and A.F. Smith. 1982. "Applications of a Method for the Efficient Computation of Posterior Distributions". *Journal of the Royal Statistical Society, Series C* 31: 214-225.

Neidhart, J., I. Szendro, and J. Krug. 2014. "Adaptation in Tunably Rugged Fitness Landscapes: The Rough Mount Fuji Model". *Genetics*, *198*(2): 699-721.

Nelson, R., and S. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press.

Nicoletti, C., and C. Rondinelli. 2010. "The (Mis)specification of Discrete Duration Models with Unobserved Heterogeneity: A Monte Carlo study". *Journal of Econometrics* 159(1): 1-13.

Norton, E., H. Wang, C. Ai. 2004. "Computing Interaction Effects and Standard Errors in Logit and Probit Models". *The Stata Journal* 4(2): 154-167.

Pavitt, K. 1984. “Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory”. *Research Policy* 13(6): 343-373.

Porter, M. 1980. *Competitive Strategy*. NY: Free Press.

Probst, G., and S.Raisch. 2005. “Organizational Crisis: The Logic of Failure”. *Academy of Management Perspectives* 19(1): 90-105.

Pyka, A. and U. Cantner. 2017. “Introduction”. In *Foundations of Economic Change: A Schumpeterian View on Behaviour, Interaction and Aggregate Outcomes*, Pyka, A. and Cantner, U. ed., 1-8, Springer.

Rivkin, J.W. 2000. “Imitation of Complex Strategies”. *Management Science* 46(6): 824-844.

Sharapov, D., P. Kattuman, V. Sena. 2011. Technological Environments, R&D Investment, and Firm Survival. MICRO-DYN Working Paper no. 34/10.

Smith, K., J. Guthrie, M. Chen. 1989. “Strategy, Size and Performance”. *Organization Studies* 10/1: 063-081.

Solanas, A., R. Selvam, J. Navarro, and D. Leiva. 2012. “Some Common Indices of Group Diversity: Upper Boundaries”. *Psychological Reports* 111(3): 777-796.

Sørensen, J.B., and T.E. Stuart. 2000. “Aging, Obsolescence, and Organizational Innovation”. *Administrative Science Quarterly* 45(1): 81-112.

Stirling, A. 2010. “Multicriteria Heterogeneity Analysis: A Novel Heuristic Framework for Apprising Energy Portfolios”. *Energy Policy* 38: 1622-1634.

Strzalecki, T. 2014. “Depth of Reasoning and Higher Order Beliefs”. *Journal of Economic Behavior and Organization* 108: 108-122.

Sutton, J. 1998. *Technology and market structure*. MIT Press, Cambridge, MA.

Teece, D. 2007. “Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance”. *Strategic Management Journal* 28: 1319-1350.

Theil, H. 1972. *Statistical Decomposition Analysis*. Amsterdam: North-Holland Publishing.

Thompson, P. 2005. “Selection and Firm Survival: Evidence from the Shipbuilding Industry, 1825–1914”. *Review of Economics and Statistics* 87(1): 26-36.

Ugur, M., E. Trushin, and E. Solomon. 2016a. “Inverted-U Relationship Between R&D Intensity and Survival: Evidence on Scale and Complementarity Effects in UK Data”. *Research Policy* 45(7): 1474–1492.

Ugur, M, E.Trushin, and E.Solomon. 2016b. “A Firm-level Dataset for Analyzing Entry, Exit, Employment and R&D Expenditures in the UK: 1997-2012”. *Data in Brief* 8: 153–157.

Verspagen, B. 2005. “Innovation and Economic Growth”, In J. Fagerberg, D.C.Mowery, and R.R.Nelson (ed.), The Oxford Handbook of Innovation, Oxford: Oxford University Press.

Vivarelli, M. 2013. “Is Entrepreneurship Necessarily Good? Microeconomic Evidence from Developed and Developing Countries”. *Industrial and Corporate Change*, *22*(6): 1453-1495.

Westhoff, F.H., B.V. Yarbrough, R.M. Yarbrough. 1996. “Complexity, Organization, and Stuart Kauffman's the Origin of Order”. *Journal of Economic Behaviour and Organization* 29: 1-25.

Williams, R. 2012. “Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects”. *The Stata Journal* 12(2): 308-331.

Wooldridge, J. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: The MIT Press.

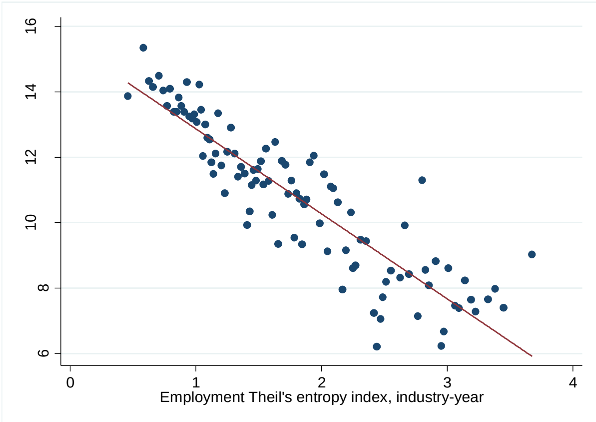
Wu, J., K.F. Lao, F. Wan, and J. Li. 2019. “Competing with Multinational Enterprises’ Entry: Search Strategy, Environmental Complexity, and Survival of Local Firms”. *International Business Review* 28: 727-738.

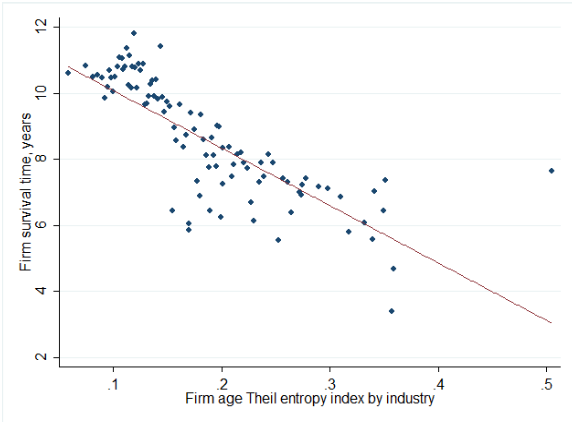
Zahra, S., and W. Bogner. 1999. “Technology Strategy and Software New Venture Performance: The Moderating Effect of Competitive Environment”. *Journal of Business Venturing* 15(2): 135-173.

Zhang, M., and P. Mohnen. 2013. “Innovation and Survival of New Firms in Chinese Manufacturing, 2000-2006”. UNU-MERIT Working Paper 2013-057. www.merit.unu.edu/publications/wppdf/2013/wp2013-057.pdf

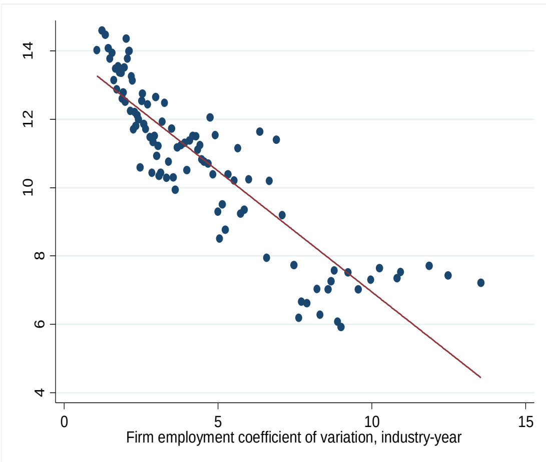
Appendix.

Panel a: Size and age heterogeneity measured with Theil index





Panel b: Size and age heterogeneity measured with coefficient of variation



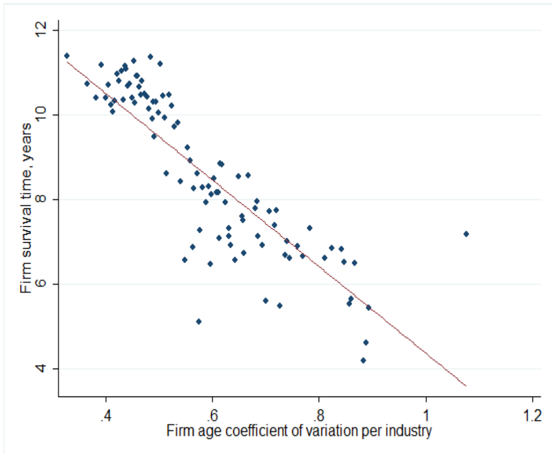
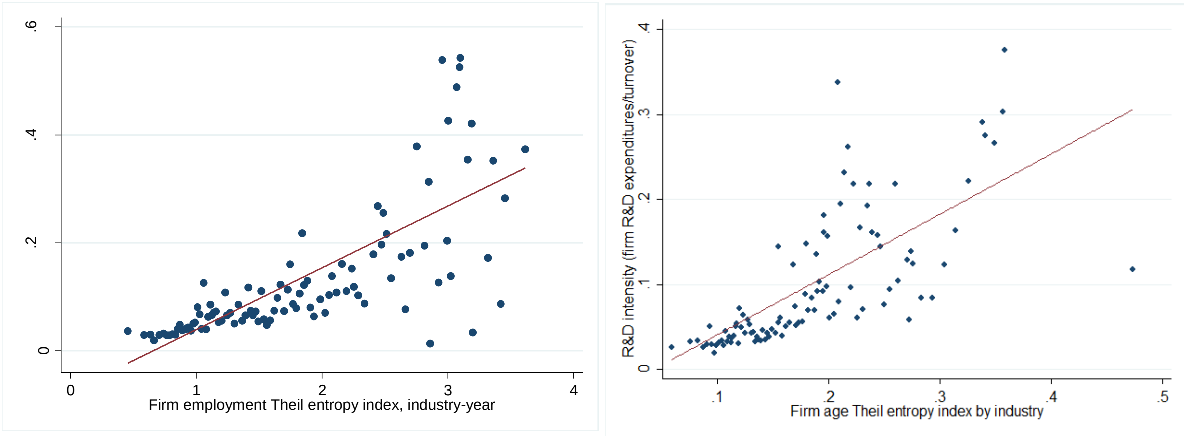


Figure 1: Survival time is declining with intra-industry size/age heterogeneity.

Panel a: Size and age heterogeneity measured with Theil index



Panel b: Size and age heterogeneity measured with coefficient of variation

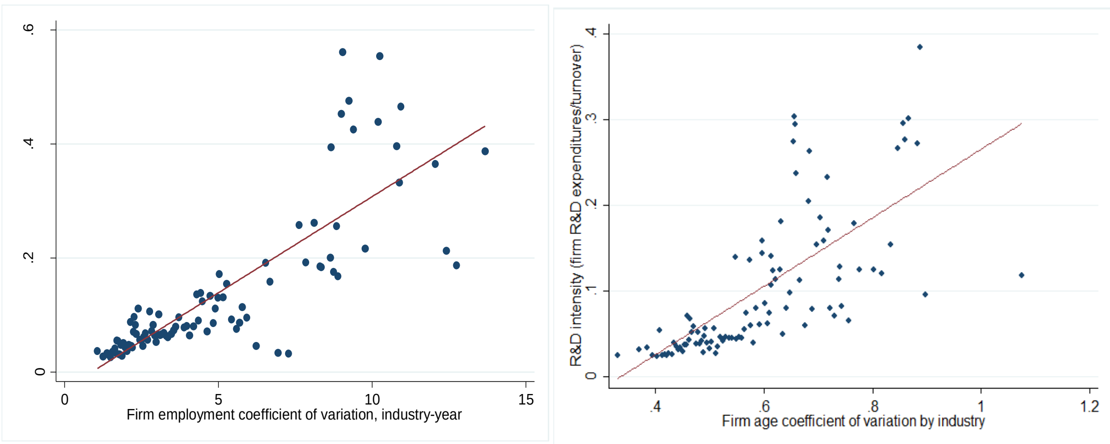
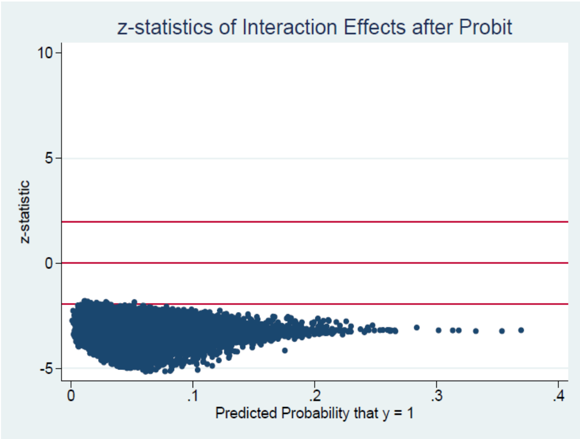
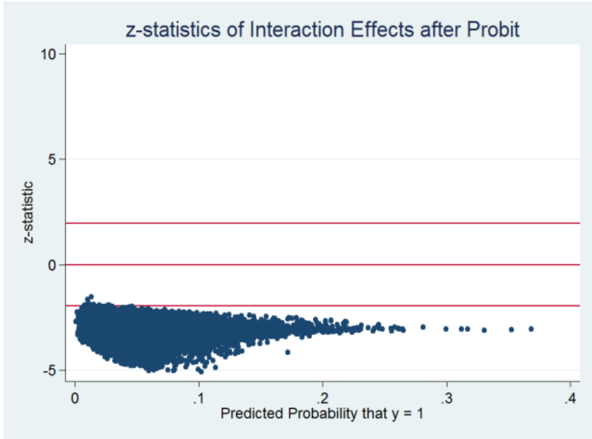


Figure 2: R&D intensity is increasing with intra-industry size/age heterogeneity.

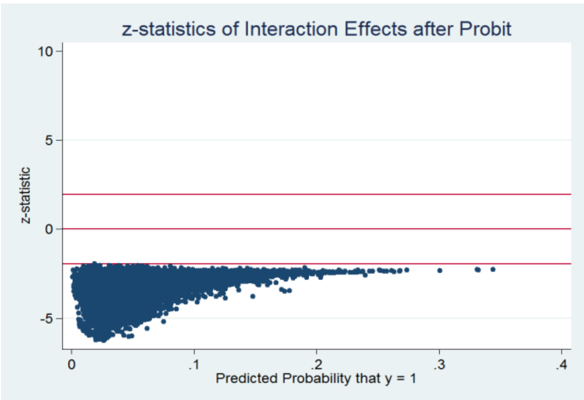
Theil entropy of firm employment.



Coefficient of variation of firm employment.



Theil entropy of firm age.



Coefficient of variation of firm age.

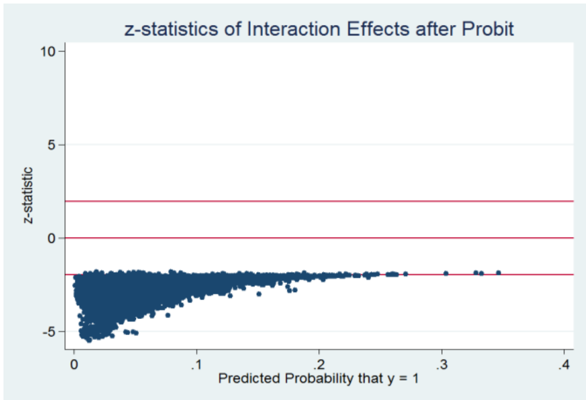


Figure 3: Interaction effect (*inteff*) plots for R&D intensity and age/size heterogeneity.

Table 1. Covariates and expected effects on exit hazard

Covariate	Description and sign of expected effect on exit hazard in brackets (+/-)	Related literature
	Covariates of main interest	
Firm-size and age heterogeneity (TI or CV)	Theil's entropy (*TI*) and coefficient of variation (*CV*) of firm employment or age in 3-digit industries (+)	Size heterogeneity is not tested for firm survival; some cohort heterogeneity is tested by Hansen and Barnett (1996) and Barnett (2008)
Interactions *TI*log(R&Dint.+1), CV*log(R&Dint.+1)*	Interaction of firm size and age heterogeneity measures with firm R&D intensity (-); natural logarithm is used	Not tested before for firm survival
	Other firm-level covariates	
R&D intensity:		
Log(R&Dint.+1)[16]	Logarithm of firm R&D intensity (-)	Aghion *et al*., (2014); Ericson and Pakes (1995)
Log(R&Dint.+1) squared	Logarithm of R&D intensity squared (+)	Aghion *et al*., (2014); Ericson and Pakes (1995); Sharapov et al., (2011); Zhang and Mohnen (2013), Ugur *et al*., (2016a).
Age: Log(firm age)	Logarithm of firm age in years (-)	Hopenhayn (1992); Ericson and Pakes (1995); Geroski (1995); Cefis and Marsili (2005); Doms *et al*., (1995); Disney *et al*., (2003)
Log(firm age) squared	Logarithm of firm age squared (+)	Agarwal and Gort (2002); Ericson and Pakes (1995); Cefis and Marsili (2005); Evans (1987)
Firm size:		
Log(Employment)	Logarithm of firm employees (-)	Hopenhayn (1992); Ericson and Pakes (1995); Geroski (1995); Cefis and Marsili (2005); Doms *et al*., (1995); Disney *et al*., (2003)
Log(Empl.) squared	Log. of firm employees squared (+)	Bhattacharjee *et al*., (2009); Cefis and Marsili (2005)
Local units: Log(Number of plants)	Logarithm of firm's local units (plants) (+)	Audretsch and Mahmood (1995); Fernandes and Paunov (2015); Audretsch (1991); Griliches and Regev (1995); Mata *et al*. (1995)
Productivity: Log(Real turnover/employees)	Logarithm of deflated turnover per employee (-)	Audretsch (1991); Hopenhayn (1992); Ericson and Pakes (1995)
Growth differential (growth_dmed)	Growth rate of firms' deflated turnover minus median industry growth rate (-)	Audretsch (1991); Hopenhayn (1992); Ericson and Pakes (1995); Cefis and Marsili (2005); Mata *et al.* (1995), Audretsch (1995), Ugur *et al.*, (2016a)
Civil R&D only	Dummy variable indicating that firm is engaged in civilian R&D only (+/ -)	Ugur *et al*., (2016a), Sharapov *et al*., (2011)
UK-owned	Dummy variable indicating that firm is UK-owned (+ /-)	Ugur *et al*., (2016a), Sharapov *et al*., (2011)
Industry covariates		

[16] The R&D to turnover ratio is augmented with one to include (scale) R&D intensity equal zero: ln(R&Dintensity=0 + 1) = 0, while the first order Taylor series approximation is $\ln(X+1) \approx X$ where $0 \leq X \leq 1$.

Concentration: Herfhindahl index	Herfindahl-Hirschman index of firm shares in industry turnover at 3-digit industry level (+/-)	McCloughan and Stone (1998); Baldwin and Rafiquzzaman (1995); Geroski (1995)
*Pavitt technology class: (Pavitt #)**	Dummy variables for Pavitt classes 1 to 5, excluded category is Pavitt 4 (+/-)	Pavitt (1984); Agarwal and Audretsch (2001); Cefis and Marsili (2005), Ugur, Trushin, Solomon (2016a)
Entry rate: Log(% entry rate)	Logarithm of firm entry rate (in %) at 3-digit SIC industry level (+)	Hannan and Freeman (1989); Fernandes and Paunov (2015)
Median industry R&D intensity: Log(Med. R&D intensity)	Logarithm of industry median ratio of total R&D to turnover at 3-digit SIC level (-)	Audretsch and Mahmood (1995); Fernandes and Paunov (2015), Ugur, Trushin, Solomon (2016a)
Firm scale in the industry: Log(Mean industry employment)	Logarithm of average employees per firm in 3-digit SIC industry level (+/-)	Fernandes and Paunov (2015); Mata and Portugal (2002); Audretsch, Houweling, and Thurik (2004)
Macroeconomic indicators		
Crisis dummy	A dummy variable equal 1 for the Asian crisis year of 1998; *dot.com* bubble crisis of 2001; and the recent financial crisis in 2008 (+)	Ugur, Trushin, Solomon (2016a); Bhattacharjee *et al*., (2009) report higher hazard rates in periods of crises
Average real effective exchange rate (Areer)	Average effective exchange rate against a basket of currencies - an increases in *Areer* indicates appreciation (+)	Bhattacharjee *et al*., (2009); Goudie and Meeks (1991)
GDP growth (%)	Annual growth rate of the GDP, % (-)	Business cycle literature; Thompson (2005) for industry output, Mata and Portugal (2002) for employment growth

Note: * Pavitt technology classes are from Pavitt (1984) as revised slightly by Bogliacino and Pianta (2010).

Table 2. Size heterogeneity and exit hazard:

Estimates from pooled and panel probit with random effects.

	Size heterogeneity is measured with Theil index		Size heterogeneity is measured with coefficient of variation	
Dependent variable: exit in year t+1	(1)	(2)	(3)	(4)
Firm Size Heterogeneity	0.0377***	0.0398***	0.0103***	0.0112***
	(0.0109)	(0.0116)	(0.0030)	(0.0032)
Size Heterogeneity*log(R&Dint.+1)	-0.244***	-0.247***	-0.0581***	-0.0588***
	(0.0588)	(0.0617)	(0.0143)	(0.0149)
Log(R&Dint.+1)	-0.578***	-0.599***	-0.755***	-0.779***
	(0.200)	(0.205)	(0.181)	(0.185)
Log(R&Dint.+1) sq.	1.250***	1.266***	1.274***	1.290***
	(0.271)	(0.282)	(0.272)	(0.283)
Log(firm Age)	-0.222***	-0.214***	-0.222***	-0.215***
	(0.0107)	(0.0114)	(0.0107)	(0.0114)
Log(Employment)	-0.197***	-0.217***	-0.197***	-0.218***
	(0.0121)	(0.0137)	(0.0121)	(0.0137)
Log(Employment) squared	0.0191***	0.0209***	0.0191***	0.0209***
	(0.0017)	(0.0019)	(0.0017)	(0.0019)
Log(Real turnover / employees)	-0.0992***	-0.104***	-0.0993***	-0.104***
	(0.0079)	(0.0079)	(0.0079)	(0.0079)
Firm growth relative to industry median growth	-0.0674***	-0.0681***	-0.0674***	-0.0681***
	(0.0102)	(0.0096)	(0.0102)	(0.0096)
Log(Number of plants)	-0.0226	-0.0202	-0.0211	-0.0186
	(0.0157)	(0.0162)	(0.0157)	(0.0162)
Civil R&D only	-0.0920***	-0.0951***	-0.0925***	-0.0957***
	(0.0130)	(0.0138)	(0.0130)	(0.0138)
UK-owned	-0.0672***	-0.0723***	-0.0665***	-0.0717***
	(0.0211)	(0.0222)	(0.0211)	(0.0222)
Log(% entry rate)	-0.0673	-0.0743	-0.0595	-0.0643
	(0.0595)	(0.0629)	(0.0599)	(0.0635)
Log(Mean industry employment)	0.0160**	0.0152*	0.0193**	0.0187**
	(0.0079)	(0.0085)	(0.0078)	(0.0083)
Log(Median R&D intensity in industry)	0.633***	0.647***	0.630***	0.642***
	(0.114)	(0.120)	(0.112)	(0.119)
Herfindahl index	-0.0434	-0.0540	-0.0387	-0.0503
	(0.0614)	(0.0651)	(0.0602)	(0.0642)
Pavitt 1	-0.0637***	-0.0578**	-0.0656***	-0.0608***
	(0.0216)	(0.0232)	(0.0220)	(0.0235)
Pavitt 2	-0.101***	-0.103***	-0.101***	-0.104***
	(0.0181)	(0.0195)	(0.0182)	(0.0195)
Pavitt 3	-0.0253	-0.0218	-0.0264	-0.0228
	(0.0234)	(0.0252)	(0.0234)	(0.0251)
Pavitt 5	-0.0494*	-0.0539**	-0.0463*	-0.0505*
	(0.0257)	(0.0274)	(0.0257)	(0.0274)
Average effective real exchange rate	0.0120***	0.0119***	0.0120***	0.0119***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Crisis dummy	0.0660***	0.0620***	0.0661***	0.0622***
	(0.0154)	(0.0159)	(0.0154)	(0.0160)
GDP growth (%)	-0.0250***	-0.0266***	-0.0251***	-0.0266***
	(0.0038)	(0.0036)	(0.0038)	(0.0036)
Constant	-1.389***	-1.379***	-1.382***	-1.374***
	(0.116)	(0.120)	(0.116)	(0.120)
$\text{Log}(\sigma_v{}^2)$		-2.351***		-2.351***
		(0.0460)		(0.0462)
N	158,316	158,316	158,313	158,313
AIC	53,695.9	53,704.9	53,692.6	53,701.1
BIC	53,935.2	53,954.2	53,931.9	53950.4

Log-likelihood	-26,824.0	-26,827.4	-26,822.3	-26,825.6
chi2	2,712.0	2,466.3	2,710.4	2,464.3
3-digit industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Correctly classified	95.66%	N/A	95.67%	N/A
p > Pearson χ^2	0.98	N/A	0.81	N/A
Area under ROC curve	0.69	N/A	0.68	N/A

Note: Firm size is natural logarithm of firm's number of employees. Top firm R&D intensity (R&Dint.) is less than 1. Estimators: pooled Probit in (1) and (3); panel random-effect Probit in (2) and (4). The dependent variable is one-year-forward exit indicator, which takes the value of 1 if firm exits in year *t+1*, and zero otherwise. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The number of firms in the sample is 35,136; the number of exiting firms is 7,802, and the number of surviving firms is 27,334. N/A: not applicable.

Table 3. Negative interaction effects for R&D intensity and age/size heterogeneity: Marginal effect estimates based on panel logit estimator with random effects.

Interaction indicator	Marginal effects at mean values by delta method	Z-stat.; standard errors are in brackets	[95% confidence interval] for the marginal effects	The range of interaction effects and their Z-statistics
Ln(R&D intensity+1)*Theil index of firm ages	-.104** (.050)	-2.949*** (.275)	[-.203; -.006]	Effect: [-.539; -.007] Z-stat.: [-6.240; -1.956]
Ln(R&D intensity+1)* Coefficient of variation of firms ages	-.047*** (.028)	-2.515*** (.2387)	[-.104; -.008]	Effect: [-.244; -.003] Z-stat.: [-5.454; -1.786]
Ln(R&D intensity)*Theil index of firm employment	-.019*** (.004)	-3.393*** (.213)	[-.028; -.010]	Effect: [-.071; -.001] Z-stat.: [-5.157; -1.789]
Ln(R&D intensity+1)* Coefficient of variation of firm employment	-.004*** (.001)	-3.240*** (.192)	[-.006; -.002]	Effect: [-.016; -.0001] Z-stat.: [-5.069; -1.535]

Note: R&D intensity is less than one. Number of observations: 158,313; number of firms: 35,136. Standard errors are in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 4: R&D intensity mitigates the hazard effects of age and size heterogeneity

Percentile of R&D intensity	R&D intensity	Average marginal effects of size heterogeneity (Theil index)	Average marginal effects of size heterogeneity (Coefficient of variation)	Average marginal effects of age heterogeneity (Theil index)	Average marginal effects of age heterogeneity (coefficient of variation)
5	.0009	.0395*** (.0116)	.0111 *** (.0032)	.9255*** (.1302)	.4819*** (.0704)
15	.0049	.0385*** (.0116)	.0108*** (.0032)	.9181*** (.1302)	.4781*** (.0704)
25	.0101	.0373*** (.0116)	.0105*** (.0032)	.9085*** (.1302)	.4733*** (.0704)
35	.0169	.0356*** (.0116)	.0101*** (.0032)	.8958*** (.1302)	.4669*** (.0704)
45	.0261	.0334*** (.0116)	.0096*** (.0032)	.8786*** (.1302)	.4582*** (.0704)
55	.0404	.0298 *** (.0116)	.0087*** (.0032)	.8518*** (.1302)	.4446*** (.0704)
65	.0656	.0236** (.0116)	.0073** (.0032)	.8047*** (.1302)	.4209*** (.0704)
70	.0859	.0186 (.0116)	.0061* (.0032)	.7668*** (.1302)	.4017*** (.0704)
75	.1122	.0121 (.0116)	.0045 (.0032)	.7176*** (.1302)	.3768*** (.0704)
80	.1467	.0036 (.0116)	.0026 (.0032)	.6532*** (.1302)	.3443*** (.0704)
85	.1927	-.0077 (.0116)	-.0001 (.0032)	.5672*** (.1302)	.3008*** (.0704)
95	.3812	-.0542*** (.0116)	-.0112*** (.0032)	.2149* (.1302)	.1227* (0704)

Note: Other covariates are taken at their mean. Standard errors are in parentheses. Top R&D intensity is less than one. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Conditional effects are based on estimated parameters using Probit random effects estimator reported in columns 2 and 4 of Table 2.